Executive Summary
Prometheus team

In this project we develop a package for analysis of experimental data from high-energy scattering experiments which is based on application of artificial neural networks (ANN).

In a high-energy scattering experiment we usually scatter a pair of protons which are building blocks of matter around us. The protons are in turn constructed from elementary particles, so called quarks and gluons. Knowing distributions of these elementary particles we can predict outcome of a high-energy scattering experiment. In other words we can predict observables which are measured in those experiments. However, a priori these distributions are not known and that is where neural networks come in handy. We can use neural networks to model the distribution of elementary particles inside the proton. We can use available experimental data to train the neural network which is the standard problem in data science.

We use a deep neural network architecture. As an input the neural network takes momentum ("speed") components of elementary particles and outputs are distribution functions which characterize probability of a particle to have given momentum components.

The neural network is a part of a more general model which contains different processing layers. The output of a model is the mean square error χ^2 which we aim to minimize.

To avoid overfitting we split the data into training and validation sets and use χ^2 of the validation set to estimate the quality of the fit. When it gets to a minimum value which doesn't change for some number of epochs we stop training the model.

To take into account errors of the experimental data we use the Monte Carlo approach. We generate a set of pseudo data (replicas). For each replica we perform a fit as was described above. The final result is obtained as an average over all replicas.

To find an optimal set of hyperparameters (number of layers, optimization algorithm etc.) we perform a separate analysis where we split the data into 4 folds. Use 3 folds to train the model and use the fourth fold to estimate the quality of the fit. The final χ^2 is obtained as an average over 4 possible choices of the fourth fold. The optimal choice of hyperparameters corresponds to the minimum value of the final χ^2.

Finally we use the optimal set of hyperparameters to train the model. After averaging over replicas we obtain a distribution function for elementary particles inside the proton which we successively use to predict experimental data.